Name: _____

GCE Stats section 6

Normal distribution

practice questions from GCE Maths part 2

Date:

Question parts in red cover content in the GCE Statistics specification but outside this section of the scheme of work

Time:

Total marks available:

Total marks achieved: _____

Questions

Q19.

A cadet fires shots at a target at distances ranging from 25 m to 90 m. The probability of hitting the target with a single shot is p. When firing from a distance d m, $p = \frac{3}{200}(90 - d)$.

Each shot is fired independently.

The cadet fires 10 shots from a distance of 40 m.

(a) (i) Find the probability that exactly 6 shots hit the target.

(ii) Find the probability that at least 8 shots hit the target.

(5)

The cadet fires 20 shots from a distance of x m.

(b) Find, to the nearest integer, the value of x if the cadet has an 80% chance of hitting the target at least once.

(4)

The cadet fires 100 shots from 25 m.

(c) Using a suitable approximation, estimate the probability that at least 95 of these shots hit the target.

(5)

(Total 14 marks)

Q20.

The number of houses sold by an estate agent follows a Poisson distribution, with a mean of 2 per week.

(a) Find the probability that in the next 4 weeks the estate agent sells,

(i) exactly 3 houses,
(ii) more than 5 houses.

(5)

The estate agent monitors sales in periods of 4 weeks.

(b) Find the probability that in the next twelve of these 4 week periods there are exactly nine periods in which more than 5 houses are sold.

(3)

The estate agent will receive a bonus if he sells more than 25 houses in the next 10 weeks.

(c) Use a suitable approximation to estimate the probability that the estate agent receives a bonus.

(6)

(Total 14 marks)

Q21.

The time taken for a randomly selected person to complete a test is M minutes, where $M \sim N(14, \sigma^2)$

Given that 10% of people take less than 12 minutes to complete the test,

(a) find the value of σ

(3)

Graham selects 15 people at random.

(b) Find the probability that fewer than 2 of these people will take less than 12 minutes to complete the test.

(3)

Jovanna takes a random sample of n people.

Using a normal approximation, the probability that fewer than 9 of these n people will take less than 12 minutes to complete the test is 0.3085 to 4 decimal places.

(c) Find the value of n.

(8)

(Total for question = 14 marks)

Q22.

In a large school, 20% of students own a touch screen laptop. A random sample of n students is chosen from the school. Using a normal approximation, the probability that more than 55 of these n students own a touch screen laptop is 0.0401 correct to 3 significant figures.

Find the value of n.

(8)

(Total for question = 8 marks)

Q23.

(a) State the conditions under which the normal distribution may be used as an approximation to the binomial distribution.

(2)

A company sells seeds and claims that 55% of its pea seeds germinate.

(b) Write down a reason why the company should not justify their claim by testing all the pea seeds they produce.

(1)

To test the company's claim, a random sample of 220 pea seeds was planted.

(c) State the hypotheses for a two-tailed test of the company's claim.

(1)

Given that 135 of the 220 pea seeds germinated,

(d) use a normal approximation to test, at the 5% level of significance, whether or not the company's claim is justified.

(7)

(Total 11 marks)

Q24.

Defects occur at random in planks of wood with a constant rate of 0.5 per 10 cm length. Jim buys a plank of length 100 cm.

(a) Find the probability that Jim's plank contains at most 3 defects.

(2)

Shivani buys 6 planks each of length 100 cm.

(b) Find the probability that fewer than 2 of Shivani's planks contain at most 3 defects.

(5)

(c) Using a suitable approximation, estimate the probability that the total number of defects on Shivani's
6 planks is less than 18.

(6)

(Total 13 marks)

Q25.

A shopkeeper knows, from past records, that 15% of customers buy an item from the display next to the till. After a refurbishment of the shop, he takes a random sample of 30 customers and finds that only
1 customer has bought an item from the display next to the till.

(a) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not there has been a change in the proportion of customers buying an item from the display next to the till.

(6)

During the refurbishment a new sandwich display was installed. Before the refurbishment 20% of customers bought sandwiches. The shopkeeper claims that the proportion of customers buying sandwiches has now increased. He selects a random sample of 120 customers and finds that 31 of them have bought sandwiches.

(b) Using a suitable approximation and stating your hypotheses clearly, test the shopkeeper's claim.
Use a 10% level of significance.

(8)

(Total 14 marks)

Q26.

A random sample of size n is to be taken from a population that is normally distributed with mean 40 and standard deviation 3. Find the minimum sample size such that the probability of the sample mean being greater than 42 is less than 5%.

(5)

(Total 5 marks)

Q27.

A woodwork teacher measures the width, w mm, of a board. The measured width, X mm, is normally distributed with mean w mm and standard deviation 0.5 mm.

(a) Find the probability that X is within 0.6 mm of w.

(2)

The same board is measured 16 times and the results are recorded.

(b) Find the probability that the mean of these results is within 0.3 mm of w.

(4)

Given that the mean of these 16 measurements is 35.6 mm,

(c) find a 98% confidence interval for w.

(4)

(Total 10 marks)

Q28.

A machine fills packets with X grams of powder where X is normally distributed with mean μ. Each packet is supposed to contain 1 kg of powder.

To comply with regulations, the weight of powder in a randomly selected packet should be such that
$P(X < \mu - 30) = 0.0005$

(a) Show that this requires the standard deviation to be 9.117 g to 3 decimal places.

(3)

A random sample of 10 packets is selected from the machine. The weight, in grams, of powder in each packet is as follows

999.8 991.6 1000.3 1006.1 1008.2 997.0 993.2 1000.0 997.1 1002.1

(b) Assuming that the standard deviation of the population is 9.117 g, test, at the 1% significance level, whether or not the machine is delivering packets with mean weight of less than 1 kg. State your hypotheses clearly.

(7)

(Total 10 marks)

Examiner's Report

Q19.

The majority of students were able to use the correct model in question (a)(i) and calculated P($X = 6$) correctly. In question (a)(ii) many students found the correct answer but not always by the quickest method which was using B(10,0.25) and finding P($X \leq 2$) . Others found P(8) + P(9) + P(10) and used the formula for B(10,0.75) or even used the formula to calculate P) ($X \leq 7$) and then subtract from 1. Some students incorrectly changed their model to Po(7.5) .

Question (b) was frequently solved efficiently and completely correctly. However, a significant minority of students did not recognise that they should take $0.2^{1/20}$ to obtain $1 - p$ instead choosing to use the tables to obtain an estimate for p , usually obtaining 0.1. Other students incorrectly thought that 0.8 (or 0.2) was the value of p and not $1 - \text{P}(X = 0)$.

Question (c) proved to be challenging. Most students were able to see that the exact distribution was B(100, 0.975) for the number of times the cadet hits the target. Few realised that they should define B(100,0.025) as the model for the number of times the cadet misses the target and then follow this with an approximation to the Poisson obtaining Po(2.5) . Even when they used Po(2.5) many were unable to change " the cadet hits at least 95 " to " the cadet misses at most 5 ". Those that stayed with B(100, 0.975) approximated to the normal, saying things like "mean = 97.5 this is too big for the tables so approximate to the normal".

Q20.

A high proportion of the candidates attempted all parts of this question successfully. Those less successful in part (a) either used Po(2) instead of Po(8), or used Po(8) but found the answer, using Poisson tables, to P($X \leq 3$) instead of P($X = 3$). A common error seen in part (a)(ii) was to write and use $\text{P}(X > 5) = 1 - \text{P}(X \leq 4)$.

In part (b) many candidates were able to find P($Y = 9$) using their answer to (a)(ii) but, for a small minority, incorrect calculations included calculating $(0.8088)^9$ or using Po(λ) with λ = 8, 9 or 24. Overall, many exemplary answers were evident for the solution to part (c) reflecting sound preparation on this topic. Marks were lost occasionally for using an incorrect, or no, continuity correction or for finding an incorrect area.

Q21.

In part (a), most students showed a good knowledge of how to standardise. However, surprisingly few were able to complete part (a) accurately and hence be awarded all 3 marks. The most common error was using a *z*-value not from the percentage points table, and hence lost the B mark due to their value being only 3.s.f.

Other errors were: adding an unnecessary continuity correction, using the variance instead of the standard deviation in their standardisation, equating to a probability rather than a *z*-value, not using the correct sign of the *z*-value and sometimes ending up on a negative standard deviation, and lastly, leaving their answer as an exact fraction.

Part (b) was well answered with the vast majority of students scoring full marks. Even if they could not find the correct model most managed to gain the method mark by writing P($X <$ 2).

Part (c) was a challenge, with very few students scoring full marks. Most students correctly identified the mean and variance for the normal approximations but many did not use the correct continuity correction.

Most students managed the correct *z*-value but a number used the wrong sign in their equation. Again, some students incorrectly used a probability as their *z*-value.

The majority of students were able to form a quadratic of some sort from their standardisation equation, and many were able to attain the M mark for a correct method of solving this. However, far too many simply produced solutions to their quadratic from no working and hence lost marks if they had used an incorrect equation.

Clearly having a quadratic in $\sqrt{n}$ was challenging, and many students had no method to deal with this – often failing to identify how it mapped to a "standard" quadratic. Some students showed good initiative and exhibited useful methods such as squaring both sides or using a substitution.

Q22.

The question was of an unusual type, requiring working backwards from an answer. Overall, candidates responded well to the initial part of the question, dealing with the normal distribution. However, algebra was then required and some surprising attempts were seen.

Many candidates appeared to be unhappy with the 'disguised' quadratic, $0.2n + 0.7\sqrt{n} - 55.5 = 0$, that is the result of correct standardisation. Some candidates recognised this as a quadratic in $\sqrt{n}$: some even wrote "let $a = \sqrt{n}$ ", to give the equation $0.2a^2 + 0.7a - 55.5 = 0$. All three of the standard methods were seen (formula, completing the square, factorisation), the first being by far the most common, and the other two very rarely.

It was not uncommon for candidates to deal with the discomfort of their $\sqrt{n}$ by rearranging and squaring. This result, if performed correctly, is a different quadratic equation: $0.04n^2 - 22.69n + 3080.25 = 0$ which nevertheless gives the same final answer of 225 (directly, rather than via 15^2). However, this alternative method involves additional work, and therefore the opportunity for things to go wrong. It was disconcerting to note that attempts to square equivalent to $(a - b)^2 = a^2 - b^2$ and $(a + b)^2 = a^2 + b^2$ were often seen. A minority of candidates who used their calculators lost marks in this question because they showed no working while solving an incorrect quadratic equation. Since the equation comes later in the question, it is perhaps unwise just to assume that your equation is correct. The only sensible option is to use a standard method, such as 'the formula', showing full details of the method used.

There was another small group of candidates who dealt with the presence of $\sqrt{n}$ in another way. These candidates had already 'standardised' correctly using mean and standard deviation of $0.2n$ and $\sqrt{0.16n}$ They dealt with the troublesome root by going back to the beginning and deliberately standardising incorrectly: this time using $0.16n$ (i.e. variance) for standard deviation. This conveniently resolves their difficulty: there is now only a linear equation in n to solve. However, they lost more marks by incorrect standardisation than they would have done by failing to solve a quadratic equation.

Q23.

Correct responses to question (a) were seen from the majority of students, although p "small" was sometimes offered. Students who stated that $np > 5, nq > 5$ usually did so in addition to, rather than instead of, the only accepted response of n is large and p is close to 0.5.

Question (b) elicited a mixed response, ranging from a null response, through "too time consuming and too expensive" to the accepted answer that there would be no pea seeds left to sell. The words "it" and "they" sometimes were used without any reference as to context.

The correct hypotheses were correctly given by a large majority of students in question (c). Sometimes they appeared in question (d), which was accepted, but occasionally different hypotheses appeared in question (c) and question (d). Occasionally either a letter other than p was used or hypotheses for a 1-tailed test were given.

In question (d) a large majority of students correctly calculated the values of the two parameters of the normal distribution, although 121 sometimes appeared as both the mean and variance following a normal approximation to the Poisson distribution Po(121).

The application of a continuity correction caused problems to many students; some were unaware that one should be used, whilst others used 135.5 instead of 134.5. Students ought to be made aware that the omission of a continuity correction resulted in the loss of at least 4 marks. A significant minority of students were able to obtain 0.0336. Few students chose to find the upper critical value. The standardisation was virtually always correctly done with the students' values.

Although nearly all students were using a 2-tailed test, it was not uncommon to see students comparing 0.0336 with 0.05 (instead of 0.025), and hence arriving at the wrong conclusion. Nevertheless, a large majority of students knew that they were expected to give a statement and contextual conclusion. The latter was sometimes marred by a failure to give a complete conclusion; in particular, the words "company's" or "seeds" were sometimes omitted.

Q24.

This question was well answered with many candidates gaining full marks. Part (b) required considerable attention to detail, and it is commendable that so many candidates were able to achieve all 5 marks. There were a small number of candidates who earned four of these five marks despite an incorrect answer to part (a). These candidates demonstrated the importance of showing one's method and including all relevant details. It is often not possible to earn 'follow through' marks unless the method and necessary detail is made explicit. The most common errors were to muddle up p and q or forget the 6 in the calculation of $P(X = 1)$.

Candidates generally gave clear, confident and accurate responses to part (c) that demonstrated mastery of both theory and detail. Although there were the usual errors of not using a continuity correction, using the wrong continuity correction or using the wrong mean and variance there were many fully correct solutions.

Q25.

This appeared to be generally a fairly straightforward question and there were many fully correct and clear solutions. In part (a) a number of candidates chose the wrong $H_1 : p < 0.15$. There are two main methods for conducting this significance test. It was noticeable that the candidates who chose the 'Critical Region' approach were generally less successful than those who evaluated the probability of obtaining an outcome 'as bad or worse' than that observed. It is advised that using the Critical Region method for hypothesis tests should be discouraged. The main errors generally involved comparing the probability found with an incorrect value. It was not uncommon to see the correct probability of 0.0480 compared with 0.05, instead of with 0.025 when a two tail test had been indicated by the hypotheses. Others incorrectly found $P(X = 1)$.

Part (b) appeared to be a lot easier than part (a) and many were able to accurately find the p value and gave good conclusions in context, Some candidates approximated to the Poisson and then to the normal so having the wrong variance for their distribution. Whilst most candidates attempted the continuity correction a substantial number wrongly used 31.5 instead of 30.5. A very small minority of the candidates attempted to use a 'Critical Region' approach however these were mostly incorrect and gained few marks and were not as successful as those who calculated a probability.

Q26.

Although many candidates clearly defined their variable there were some who wrote $X \sim N\left(40, \frac{9}{n}\right)$ but clearly meant $\bar{X}$ as their standardization indicated. Most used a z value of 1.6449 but a number used an incorrect inequality and ended up with $n < 6.087$...and gave a final answer of 6.

Q27.

The first two parts of this question proved quite challenging for many candidates who were unable to handle the modulus. Many simply missed the significance of the phrase "...is within..." and others misinterpreted the ranges dividing by 2. Those who did interpret the question correctly usually had few problems in calculating the required probabilities although there were some errors with the variance in part (b).

The confidence interval in part (c) was answered very well by most candidates and many correct solutions were seen. Few failed to use z = 2.3263 and only a small minority used an incorrect standard error.

Q28.

Marks were lost by some students in question (a) by failing to have consistency in the use of signs. The test was well done by all but the weaker students who usually failed to use the correct standard error. However the incorrect use of signs too often lost a mark in the actual test, but these students usually recovered to come to a correct conclusion. Most students gave the result well in context but some lost a mark by not using the key words and /or not concentrating on the "less than" idea. Standardising correctly also proved to be an issue here, but many students scored well on this part.

Mark Scheme

Section 6	Normal distribution part 2
qu.19	June 2014 S2 - Q4c
qu.20	May 2012 S2 - Q4c
qu.21	June 2017 S2 - Q5ac
qu.22	June 2016 S2 - Q5
qu.23	June 2014 S2 - Q5a
qu.24	May 2011 S2 - Q5c
qu.25	May 2011 S2 - Q6b
qu.26	June 2013 S3R - Q2
qu.27	June 2010 S3 - Q3
qu.28	May 2014 S3 - Q7a

Q1.

Question Number	Scheme	Marks
(a)	X is the random variable the Number of successes, $X \sim B(10, 0.75)$	B1
(i)	$P(X = 6) = (0.75)^6 (0.25)^4 \, {}^{10}C_6$ or $P(X \le 6) - P(X \le 5)$	M1
	$= 0.145998$ awrt 0.146	A1
(ii)	Using $X \sim B(10, 0.75)$	
	$P(X \ge 8) = P(X = 8) + P(X = 9) + P(X = 10)$	M1
	$= (0.75)^8 (0.25)^2 \, {}^{10}C_8 + (0.75)^9 (0.25)^1 \, {}^{10}C_9 + (0.75)^{10}$	
	$= 0.52559$ awrt 0.526	A1
	Or	
	Using $Y \sim B(10, 0.25)$ and $P(Y \le 2) = 0.5256$	(5)
(b)	$1 - P(0) = 0.8$ or $P(0) = 0.2$	M1
	$(1-p)^{20} = 0.2$	
	$1 - p = 0.9227$	
	$p = 0.0773$	A1
	$\frac{3}{200}(90 - x) = 0.0773$	M1
	$x = 84.84$	
	$x = 85$	A1cao (4)
(c)	X – successes ~B(100, 0.975)	B1
	Y – not successes ~B(100, 0.025)	
	Y~Po(2.5)	M1A1
	$P(Y \le 5) = 0.958$	M1A1 (5)

	Notes	[14]
(a)	B1 writing or using $p = 0.75$ or $p = 0.25$ anywhere in (a)(i) or (a)(ii)	
(i)	M1 writing or using $(p)^6(1-p)^4\ {}^{10}C_6$ **or** writing for p =0.75, $P(X \le 6) - (X \le 5)$ **or** for $p = 0.25$, $P(X \le 4) - P(X \le 3)$ **or** correct answer.	
(ii)	M1 writing B(10, 0.75) and writing or using $P(X=8)+P(X=9)+P(X=10)$ oe **or** writing B(10, 0.25) and writing or using $P(Y \le 2)$.	
	Using correct Binomial must be shown by $(0.75)^n(0.25)^{10-n}$ or a correct answer.	
(b)	M1 for writing or using $1 - P(0) = 0.8$ or $P(0) = 0.2$ or $(1-p)^{20} = 0.2$. Allow any inequality sign.	
	A1 awrt 0.0773 or awrt 0.923.	
	M1 subst in $\frac{3}{200}(90-x)$ for p **NB** this may be substituted in earlier for p.	
	Allow for $\frac{3}{200}(90-x) = k$ where $0 < k < 1$ $k \neq 0.8$ or 0.2 Allow any inequality sign	
	A1 condone $x \ge 85$. Do not allow $x \le 85$.	
(c)	B1 writing or using 0.975 or 0.025, may be implied by Po(2.5)	
	M1 using Po approximation	
	A1 Po(2.5)	
	M1 writing or using $P(Y \le 5)$	
	A1 awrt 0.958	
	SC use of normal approximation can get B1 M0A0M1A0	
	B1 writing or using 0.975 or 0.025 implied by normal with mean 97.5 or answer of 0.973	
	M1 for awrt 0.973	

Q20.

Question Number	Scheme		Marks
(a)	Let X be the random variable the number of houses sold.		
	X ~Po(8)		B1
(i)	$P(X \leq 3) - P(X \leq 2) = 0.0424 - 0.0138$ or $\frac{e^{-8}8^3}{3!}$		M1
	$= 0.0286$	awrt 0.0286	A1
(ii)	$P(X > 5) = 1 - P(X \leq 5)$		M1
	$= 1 - 0.1912$		
	$= 0.8088$	awrt 0.809	A1
			(5)
(b)	Let Y be the random variable = the number of periods where more than 5 houses are sold		
	Y~ B(12,0.8088)		M1
	$P(Y = 9) = (0.8088)^9(1 - 0.8088)^3\frac{12!}{9!3!}$		M1
	$= 0.228$	awrt 0.228	A1
			(3)
(c)	N(20,20)		M1A1
	$P(X > 25) = 1 - P\left(Z \leq \frac{25.5 - 20}{\sqrt{20}}\right)$		M1,M1,A1
	$= 1 - P(Z \leq 1.23)$		
	$= 1 - 0.8907$		
	$= 0.1093 / 0.1094$	awrt 0.109	A1
			(6)
			Total 14

Notes

(a) 1st B1 for writing or using Po(8) in either (i) or (ii)

(i) M1 writing or using $P(X \leq 3) - P(X \leq 2)$ or $\frac{e^{-8}8^3}{3!}$

(ii) M1 writing or using $1 - P(X \leq 5)$

(b) M1 writing or attempting to use B(12,their (a(ii))) NB ft their a(ii) to at least 2sf

M1 $\frac{12!}{9!3!}$ $(a(ii))^9(1 - a(ii))^3$ allow ${}^{12}C_3$ or ${}^{12}C_9$ or 220 instead of $\frac{12!}{9!3!}$ NB ft their a(ii) to at least 1sf but an expression must be seen (No use of tables)

(c) 1st M1 for writing or using a normal approximation

1st A1 for correct mean and sd (may be given if correct in standardisation formula)

2nd M1 Standardising using their mean and their sd and using [24.5, 25, 25.5, 26 or 26.5] and for finding correct area by doing 1 – P(Z < "their 1.23")

NB if they have not written down a mean and sd then they need to be correct in the standardisation to gain this mark.

3rd M1 for attempting a continuity correction (26 ± 0.5)

2nd A1 for $\pm\frac{25.5 - 20}{\sqrt{20}}$ or ± awrt 1.2 or better.

SC using P(X< 26.5/25.5) – P(X<25.5/24.5) can get M1A1 M0M1A0A0

Q21.

Question Number	Scheme		Marks
(a)	$P(M<10)=P\left(Z<\frac{12-14}{\sigma}\right)=0.1$		
	$\Rightarrow \frac{12-14}{\sigma}=,\ -1.2816$	**M1** standardising ($\pm$) with 12, 14 and σ and setting equal to a z value where $\lvert z\rvert>1$ **B1** ± 1.2816 or better	**M1** **B1**
	$\sigma=1.5605\ldots\ldots$ =awrt 1.56 minutes	**A1** awrt 1.56 Do **not** allow answer written as an exact fraction.	**A1** (3)
(b)	T represents number less than 12 minutes. $T\sim B(15, 0.1)$	**B1** Writing or using B(15, 0.1) .	**B1**
	$P(T\le 1)$	**M1** writing $P(T\le 1)$ or $P(T<2)$ any letter may be used.	**M1**
	$=0.549$	**A1** awrt 0.549	**A1**
		NB 0.549 gets B1 M1 A1	(3)
(c)	[$T\sim$ number of people who take less than 12 mins to complete the test] $T\sim B(n, 0.1)$		
	T can be approximated by $N(0.1n, 0.09n)$	**B1** mean = $0.1n$ and Var = $0.09n$ **oe** may be seen in an attempt at standardisation	**B1**
	$P\left(Z<\frac{8.5-0.1n}{\sqrt{0.09n}}\right)=0.3085$	**M1** using a continuity correction either 8.5 or 7.5 in an attempt at standardised form. Allow 0.09 for sd.	**M1**
	$\frac{8.5-0.1n}{\sqrt{0.09n}}=-0.5$ or $\frac{8.5-0.1x^2}{0.3x}=-0.5$	**B1** a z value of awrt $\pm$ 0.5	**B1**
		M1 standardising using their mean and sd. (If these have not been given then they must be correct here) **and** one of 7.5, 8, 8.5, 9 or 9.5 **and** equal to a z value where $\lvert z\rvert>0.4$. Allow any form	**M1**
		A1 A correct equation in **any form**. ISW. Do **not** allow if they have $0.3n$ rather than $0.3\sqrt{n}$	**A1**
	$0.1n-0.15\sqrt{n}-8.5=0$ $\sqrt{n}=10$	**M1** using either the quadratic formula or completing the square or factorising or any correct method to solve **their 3 term quadratic**. If they write the quadratic formula down then allow one slip. If no formula written down then it must be correct for their equation. May be implied by seeing 10 or 8.5. They must show working if the equation used is not correct. **2nd A1** awrt 10.0 – do not need to see n or $\sqrt{n}$. Allow $n=10$ May be implied by 100	**M1A1**
	$n=100$	**3rd A1 cso** 100 If they have a second answer of 72.25 they must reject it to get this final mark.	**A1cso** (8)
			(Total 14)

Q22.

Question Number	Scheme		Marks
	$N(0.2n, 0.16n)$	B1: Mean $= 0.2n$ and Var $= 0.16n$ oe this may be awarded if they appear in the standardisation as $0.2n$ and either $0.16n$ or $\sqrt{0.16n}$	B1
	$P\left(Z > \frac{55.5-0.2n}{\sqrt{0.16n}}\right) = 0.0401$	M1: Using a continuity correction either 55.5 or 54.5	M1
	$\frac{55.5-0.2n}{\sqrt{0.16n}} = 1.75$	B1: Using a z = awrt $\pm$ 1.75 M1: Standardising using either 55.5, 54.5 or 55 and equal to a z value. Follow through their mean and variance. If they have not given the mean and Var earlier then they must be correct A1: A correct equation. May be awarded for $\frac{55.5-0.2n}{\sqrt{0.16n}} = 1.75$ Condone use of an inequality sign rather than an equals sign	B1M1A1
	$0.2n + 0.7\sqrt{n} - 55.5 = 0$	M1**d**: This is dependent on the previous method mark being awarded. Using either the quadratic formula or completing the square or factorising or any correct method to solve **their 3 term** equation. If they write the formula down then allow a slip. If no formula written down then it must be correct for their equation. May be implied by correct answer or $\sqrt{n} = 15$ or 342.25 **NB** you may award this mark if they use 54.5 for awrt 14.9, -18.4, 221 or 337 55 for awrt -18.4, 14.9,223 or -117 If the answer is not one of these then the method for solving their 3 term equation must be seen.	M1**d**
	$\sqrt{n} = 15$	A1: Allow 15 or -18.5 do not need to see n or $\sqrt{n}$. Condone $n = 15$ or $n = -18.5$	A1
	$n = 225$	A1 : cao 225 do not need to see n or $\sqrt{n}$	A1 (8)

	Alternative method for last 3 marks $(0.2n-55.5)^2=(-0.7\sqrt{n})^2$ $0.04n^2-22.69n+3080.25=0$ $n=225$ or 1369/4 $n=225$	M1 solving 3 term quadratic in n as above A1 either 225 or 1369/4 or 342.25 A1must select 225	**Total 8**

Q23.

Question Number	Scheme	Marks
(a)	n is large and p close to 0.5	B1B1 (2)
(b)	There would be no pea seeds left	B1 (1)
(c)	H_0: $p = 0.55$ H_1 : $p \neq 0.55$	B1 (1)
(d)	X~N(121, 54.45)	B1
	$P(X \geq 134.5) = P\left(Z \geq \frac{134.5-121}{\sqrt{54.45}}\right)$ or $\pm\frac{x-0.5-121}{\sqrt{54.45}} = 1.96$	M1M1A1
	$= P(Z \geq 1.8295..)$	
	$= 1 - 0.9664$	
	$= 0.0336/0.0337$ $\quad x = 135.96$	A1
	Accept H_0 not in CR, not significant	M1
	The **company's claim** is justified **or** **55%** of its pea **seeds germinate**	A1cso (7)
	Alternative	
	X~N(99, 54.45)	B1
	$P(X \leq 85) = P\left(Z \leq \frac{85.5-99}{\sqrt{54.45}}\right)$ or $\pm\frac{x+0.5-99}{\sqrt{54.45}} = 1.96$	M1 M1 A1
	$= P(Z \geq 1.8295..)$	
	$= 1 - 0.9664$	
	$= 0.0336/0.0337$ $\quad x = 107.5$	
	Accept H_0 not in CR, not significant	M1
	The **company's claim** is justified **or** **55%** of its pea **seeds germinate**	A1cso [11]
	Notes	
(a)	B1 accept $n > 50$ (or any number bigger than 50) B1 p close to 0.5 NB Do not accept $np > 5$, $nq > 5$.	
(b)	Must have the idea of no peas left. They must mention either **pea** or **seeds**.	
(c)	B1 both hypotheses correct. Must use p or π and 0.55 oe. Accept the hypotheses in part (d).	
(d)	B1 correct mean and Var, may be seen in the standardiation formula as 121 and $\sqrt{54.45}$ or 7.38 to 2dp or implied by a correct answer M1 for attempting a continuity correction (Method 1:135/85 ± 0.5 / Method 2:x ± 0.5) M1 for standardising using their mean and their standard deviation and using either Method 1 [134.5, 135, 135.5, 85, 85.5 or 84.5 accept ± z.] Method 2 [(x ± 0.5) and equal to a ± z value] A1 correct z value awrt ± 1.83 **or** $\pm\frac{134.5-121}{\sqrt{54.45}}$ $\left(\frac{85.5-99}{\sqrt{54.45}}\right)$ or $\pm\frac{x-0.5-121}{\sqrt{54.45}} = 1.96$ $\left(\pm\frac{x+0.5-99}{\sqrt{54.45}} = 1.96\right)$ or(allow 1.6449 if 1 tail test in (c)) A1 awrt 0.0336/0.0337 **or** awrt 136 (allow 126 if one tail test in (c)) or a comparison of awrt1.83 with 1.96 (1.6449) M1 A correct statement. Accept H_0, oe if a 2-tailed test in (c), reject H_0, oe if a 1-tailed test in (c). Allow for a correct contextual statement. Do not allow contradictions of non-contextual statements. A1 A correct contextual statement to include words in bold/underlined for a 2-tailed test. This is not a follow through mark. **NB** if finding $P(X=135)$ they can get B1 M1 M1 A0 A0 M0 A0	

Q24.

Question Number	Scheme	Marks
(a)	$X \sim \text{Po}(5)$; $\text{P}(X \leq 3) = 0.2650$	M1 A1 (2)
(b)	Let Y = the no.of planks with at most 3 defects, $Y \sim$Binomial $Y \sim \text{B}(6, 0.265)$ $\text{P}(Y < 2) = \text{P}(Y \leq 1)$ $= \left[0.735^6 + 6 \times 0.265 \times 0.735^5\right]$ $= 0.4987\ldots$ **awrt 0.499 or 0.498**	M1 A1ft M1 A1 A1 (5)
(c)	Let T = total number of defects on 6 planks, $T \sim \text{Po}(30)$ so $T \approx S \sim$Normal $S \sim \text{N}(30, 30)$ $\text{P}(T < 18) = \text{P}(S < 17.5)$ $= P\left(z < \frac{17.5 - 30}{\sqrt{30}}\right)$ $= \text{P}(Z < -2.28\ldots)$ $= 0.01123\ldots$ **awrt 0.0112 or 0.0113**	M1 A1 M1 M1 A1 A1 (6) **13**

Notes:

(a)
M1 for identifying Po(5) - it should be clearly seen somewhere or implied
A1 for correct probability. Allow 0.265

(b)
1st M1 for writing or using the binomial - may be implied by use of $nq^x(1-q)^{6-x}$ with $n \geq 1$
1st A1ft for $n = 6$ and p = their (a) may be implied by $6p(1-p)^5$ or $(1-p)^6$
NB if they write B(6,(a)) they get M1 A1
2nd M1 for writing $\text{P}(Y \leq 1)$ **or** $\text{P}(Y = 0) + \text{P}(Y = 1)$ **or** $(1-q)^6 + nq(1-q)^5$ with $n \geq 1$
2nd A1 $(1-p)^6 + 6p(1-p)^5$ where p = their (a)
3rd A1 for awrt 0.499
SC use of a probability in the tables – lose last two marks – could get M1A1M1 M0 A0

(c)
1st M1 for a normal approx
1st A1 for correct mean and sd
2nd M1 for use of continuity correction, either 17.5 or 18.5 or 42.5 or 41.5 seen
3rd M1 Standardising with their mean and their sd and 17.5 or 18 or 18.5 or 41.5 or 42 or 42.5
NB if they have not written down a mean and sd then they need to be correct in the standardisation to gain this mark.
2nd A1 for $z = \pm 2.28$ or better. May be awarded for $\pm\frac{17.5 - 30}{\sqrt{30}}$ [NB no continuity correction $z = 2.19$]
3rd A1 for awrt 0.0112 or 0.0113 [NB no approximation gives 0.00727…]
SC using P(X<18.5) – P(X<17.5) can get M1 A1 M1 M0A0A0

Q25.

Question Number	Scheme	Marks
(a)	$H_0 : p = 0.15$ $H_1 : p \neq 0.15$	B1 B1
	$X \sim B(30, 0.15)$	M1
	$P(X \leq 1) = 0.0480$ or CR: $X = 0$	A1
	$(0.0480 > 0.025)$	
	not a significant result or do not reject H_0 or not in CR	M1
	there is no evidence of a change in the proportion of customers buying an item from the display.	A1ft
		(6)
(b)	$H_0 : p = 0.2$ $H_1 : p > 0.2$	B1
	Let S = the number who buy sandwiches, $S \sim B(120, 0.2)$,	
	$S \approx W \sim N\left(24, \sqrt{19.2}^2\right)$	M1 A1
	$P(S \geq 31) = P(W \geq 30.5)$	M1
	$= P\left(Z > \frac{30.5 - 24}{\sqrt{19.2}}\right)$ or $\frac{x - 0.5 - 24}{\sqrt{19.2}} = 1.2816$	M1
	$[= P(Z > 1.48..)]$	
	$= 1 - 0.9306$	M1
	$= 0.0694$ $x = 30.1$	A1
	< 0.10 so a significant result, there is evidence that more customers are purchasing sandwiches or the shopkeepers claim is correct.	B1ft
		(8)
		14

Notes:

(a)

1st B1 for H_0 must use p 2nd B1 for H_1 must use p

1st M1 for writing or using B(30,0.15) – may be implied by correct CR

1st A1 0.0480 or $X = 0$. Allow $X \leq 0$. Ignore upper CR. NB Allow CR $X \leq 1$ if using one tail test.

2nd M1 A correct statement (see table below) Do not allow non-contextual conflicting statements eg"significant" and "accept H_0". **Ignore comparisons**

2nd A1 for a correct statement in context. For context we need idea of change/decrease in number of customers buying from display – may use different words. NB A correct contextual statement on its own scores M1A1

	Two tail $0.025 < p < 0.975$ **or** **One tail** $0.05 < p < 0.95$	**Two tail** $p < 0.025$ or $p > 0.975$ **or** **One tail** $p < 0.05$ or $p > 0.95$
2nd M1	not significant/ accept H_0/ Not in CR or contextual	significant/ reject H_0/ In CR or contextual
2nd A1	There is no evidence of a change/decrease in the proportion of customers buying an item from the display	There is evidence of a change/decrease in the proportion of customers buying an item from the display.

(b)

1st B1 both hypotheses correct – must use p.

1st M1 for a normal approx

1st A1 for correct mean and sd

2nd M1 for use of continuity correction, either 30.5 or 31.5 or $(x \pm 0.5)$ seen

3rd M1 standardising with their mean and their sd and 30.5, 31 or 31.5 or x or $(x \pm 0.5)$)

4th M1 for 1 - tables value or 1.2816

2nd A1 for awrt 0.069 or $x = 30.1$

2nd B1ft For a correct conclusion in context using their probability and 0.1 For context we need idea of more customers buying sandwiches – may use different words

	One tail $0.1 < p < 0.9$ **or Two tail** $0.05 < p < 0.95$	**One tail** $p < 0.1$ or $p > 0.9$ **or Two tail** $p < 0.05$ or $p > 0.95$
2nd M1	not significant/ accept H_0/ Not in CR or contextual	significant/ reject H_0/ In CR or contextual
2nd A1	There is no evidence of an increase in the proportion of customers buying sandwiches	There is evidence of a change/increase in the proportion of customers buying sandwiches.

Q26.

Question Number	Scheme	Marks
	$X \sim N(40, 3^2)$ $\quad \overline{X} \sim N\left(40, \frac{9}{n}\right)$ $\qquad$ (Condone $Y \sim N\left(40, \frac{9}{n}\right)$)	B1
	$P(\overline{X} > 42) = P\left(Z > \frac{42-40}{\sqrt{\frac{9}{n}}}\right)$	M1
	$\frac{42-40}{\sqrt{\frac{9}{n}}} \geq 1.6449$	B1 dM1
	$n \geq 6.087$ $n = 7$	A1 **[Total 5]**

1^{st} B1 for stating the correct distribution for $\overline{X}$.
May be implied if correctly used in line 2 and no incorrect version seen elsewhere.

1^{st} M1 for an attempt to standardise with 42, 40 and their $\sqrt{\frac{9}{n}}$, must have n. Allow $\pm$

2^{nd} B1 for using $z = \pm 1.6449$ (or better)

2^{nd} dM1 Dep on 1^{st} M1 for forming an equation in n or $\sqrt{n}$. Allow "=" or "<"
i.e. setting their standardised expression = their z value ($|z| > 1.5$)

A1 for $n = 7$ only
The A1 must follow from correct working so e.g. $n < 6.087$ leading to $n = 7$ is A0

Q27.

Question Number	Scheme	Marks
(a)	$E \sim \mathrm{N}(0, 0.5^2)$ or $X \sim \mathrm{N}(w, 0.5^2)$	
	$\mathrm{P}(\lvert E\rvert < 0.6) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.6}{0.5}\right)$ or $\mathrm{P}(\lvert X - w\rvert < 0.6) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.6}{0.5}\right)$	M1
	$= \mathrm{P}(\lvert Z\rvert < 1.2)$	
	$= 2 \times 0.8849 - 1 = 0.7698$ awrt **0.770**	A1
		(2)
(b)	$\overline{E} \sim \mathrm{N}\left(0, \frac{1}{64}\right)$ or $\overline{X} \sim \mathrm{N}\left(w, \frac{0.5^2}{16}\right)$	M1
	$\mathrm{P}(\lvert \overline{E}\rvert < 0.3) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.3}{\frac{1}{8}}\right)$ or $\mathrm{P}(\lvert \overline{X} - w\rvert < 0.3) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.3}{\frac{1}{8}}\right)$	M1, A1
	$= \mathrm{P}(\lvert Z\rvert < 2.4)$	
	$= 2 \times 0.9918 - 1 = 0.9836$ awrt **0.984**	A1
		(4)
(c)	$35.6 \pm 2.3263 \times \frac{1}{8}$	M1 B1
	(**35.3, 35.9**)	A1,A1
		(4)
		10

(a) 1st M1 for identifying a correct probability (they must have the 0.6) and attempting to standardise. Need | |. This mark can be given for 0.8849 - 0.1151 seen as final answer.
1st A1 for awrt 0.770. NB an answer of 0.3849 or 0.8849 scores M0A0 (since it implies no | |)
M1 may be implied by a correct answer

(b) 1st M1 for a correct attempt to define $\overline{E}$ or $\overline{X}$ but must attempt $\frac{\sigma^2}{n}$. Condone labelling as E or X
This mark may be implied by standardisation in the next line.
2nd M1 for identifying a correct probability statement using $\overline{E}$ or $\overline{X}$. Must have 0.3 and | |
1st A1 for correct standardisation as printed or better
2nd A1 for awrt 0.984
The M marks may be implied by a correct answer.

Sum of 16, not means
1st M1 for correct attempt at suitable sum distribution with correct variance ($= 16 \times \frac{1}{4}$)
2nd M1 for identifying a correct probability. Must have 4.8 and | |
1st A1 for correct standardisation i.e. need to see $\frac{4.8}{\sqrt{4}}$ or better

(c) M1 for $35.6 \pm z \times \frac{0.5}{\sqrt{16}}$
B1 for 2.3263 or better. Use of 2.33 will lose this mark but can still score ¾
1st A1 for awrt 35.3
2nd A1 for awrt 35.9

Q28.

Question Number	Scheme	Marks
(a)	$z = \pm 3.2905$	B1
	$\sigma = \frac{30}{3.2905}$	M1
	$\sigma = 9.117$ **	A1cso
		(3)
(b)	$H_0: \mu = 1000$ $H_1: \mu < 1000$	B1
	mean weight = 999.54	B1
	$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{(999.54-1000)}{\frac{9.117}{\sqrt{10}}} = -0.160$ or $\frac{c-1000}{\sqrt{83.12/10}} = -2.3263 \therefore$ CR $c < 993.29$	M1A1
	1% critical value = -2.3263	B1
	$-2.3263 < -0.160$	
	Accept H_0 / not in critical region	dM1
	There is no evidence that that the machine is delivering packets of mean weight less than 1 kg	A1ft
		(7)
		Total 10

Notes	
(a)	M1 for 30/'their $\lvert z \rvert$ ' >1
	A1 cso as given answer
(b)	1st B1 both hypotheses correct.
	Accept 1kg in hypotheses if consistent units used in working usually either kg or g.
	2nd B1 999.54 (g) or 0.99954 (kg)
	1st M1 for standardising using their mean allow $\pm$, 1000 and $\frac{9.117}{\sqrt{10}}$ o.e. in kg
	1st A1 awrt -0.160 unless clearly using $\lvert z \rvert$ (stated) then accept 0.160 or CR awrt 993
	Condone -0.16 if fully correct expression seen.
	3rd B1 $\pm$ 2.3263 sign consistent with test statistic or $p = 0.4364 > 0.01$ NB $p = 0.5636 < 0.99$
	2nd dM1 dependent upon 1st M for a correct statement linking their test statistic and their cv
	Contradictory statements score M0 e.g. "significant, do not reject H_0"
	2nd A1 for correct conclusion in context. Must mention 'machine' and 'packets'.